第 28 卷第 1 期　　数学的实践与认识　　Vol. 28 No. 1
1998 年 1 月　　MATHEMATICS IN PRACTICE AND THEORY　　Jan., 1998

零件的参数设计

叶庆春　侯小虎　刘淑君

指导教师：　甘湘华

(四川联合大学 (四川大学)，成都 610064)

编者按　本文针对参数最优设计问题的实际背景提出综合考虑损失函数与元件成本的概率统计最优化模型，并用模拟褪火法寻求最优解，较好地解决了最优参数设计问题，论文还用统计实验及泰劳展开论证了参数 y 近似服从正态分布的结论．论文叙述简洁、语言流畅、有一定的参考价值．

摘　要　本文解决解决的问题是一个零件参数优化设计问题．

首先我们利用统计抽样方法给出了频率直方图，估计参数 y 应当服从正态分布，进一步利用经验公式作多元一阶 Taylor 展开，得到一个正态随机变量的线性组合式，确定 y 近似正态分布，且用 χ^2 拟合优度加以检验．然后可直接求出原设计的生产总费用为 3,074,790 元，接着对第地二个问题建立了一个优化模型，用模拟退火全局优化算法，得到第二个问题的解：标定值依次为 0.0829,0.3223,0.1195,0.0881,1.8097 ,12.9960,0.6297，总费用合为 434,681.2352 元，与原设计相比总费用减少：2, 639, 109.7648 元。

最后，我们用统计抽样进行模型的检验，以及给出了模型的评价与改进方向．

一、问题的重述与背景分析 (略)

二、模型的基本假设及符号说明

1. 假设零件参数的标定值代表期望值，在生产部门无特殊要求时，容差为均方差的 3 倍；

2. 假设粒子分离器参数 y 偏离 $y_0\pm0.1$ 时，产品为次品，质量损失为 1000 元；

当 y 偏离参数 $y_0=\pm0.3$ 时，产品为废品，损失为 9000 元；

3. 零件参数的标定值有一定的容许范围；

4. 容差分为 A,B,C 三个等级，用与标定值的相对值表示，A 等为 $\pm1\%$,B 等为 $\pm5\%$,C 等为 $\pm10\%$；

5. 假设 7 个零件随机误差都符合正态分布．

文中所用符号说明

$A=(a_{ij})_{7\times3}$	成本矩阵;	D_i	第 i 种零件的容差行向量;
$B=(b_{ij})_{7\times2}$	容差设计矩阵;	X_2	标定值的上界;
y_0	y 的目标值;	δ_i	x_i 的均方差;
δ	y的均方差;	X_i	标定值的下界;
B_i	第 i 行向量;	C_1	一件产品的生产成本;
P_2	产品是废品的概率;	$\bar{y}$	y 的期望值;
P_1	产品是次品的概率;	C_2	一件产品的质量损失的期望值;
value	生产总费用;		

$X=(x_1,x_2,\cdots,x_n)$　标定值的向量；　$f(x_1,\cdots,x_n)$　经验函数；

$$b_{ij}=\begin{cases}1 & \text{当 } i \text{ 的容差为 } j \text{ 时,}\\ 0 & \text{否则.}\end{cases}$$

三、模型的分析与建立

(一) 确定 y 的分布函数

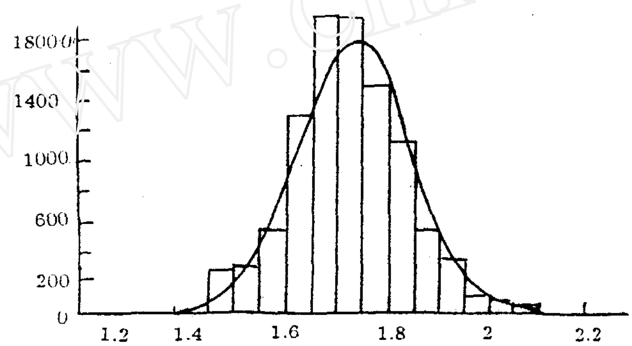
a) 利用统计抽样来估计 y 的分布

由经验公式不能直接看出参数 y 的分布，于是为了对 y 的分布有一个直观上的了解，我们采用统计抽样的方法绘出 y 的频率直方图.

直方图的画法： 为随机数落入各区间 $[t_{i-j},t_i]$ 内的个数， $f_i=\frac{v_i}{n}$.

在 xy 平面上对每一个 i, 以 $[t_{i-j},t_i]$ 为底，$y_i=\frac{f_i}{\Delta}$ 为高画小长形即为直方图，容易看出，样本容量越大，分组越细，则直方图越接近密度函数曲线下的曲边梯形，数学软件 Maplev 3 统计包中画直方图即根据上面原理.

为了得到较直观、准确的直方图，取了 500 个样本点即 y 值，直方图如下：



由图可以估计， y 的分布近似于正态分布.

b) 受 (a) 的启发，我们便想从经验公式出发推导出 y 的概率密度函数.

方法 1 从理论上严格推导 y 的概率密度函数. 我们实践后发现，既使利用概率密度函数的和、差、商能算出 y 的发布，其表达式也是相当复杂且得不到一个确定的表达式，因此，此方法在理论上可行，在实际中则要视具体情况而定.

方法 2 由经验方程及 x_i 所容许的范围，我们知道 y 对 x_i 任意的 n 阶偏导连续则可将其在标定值 $Z_0=(x_{i0})_{i=1\cdots 7}$ 作 Taylor 展开式：

$$y=f(x_{10},x_{20},\cdots,x_{70})+\frac{\partial f}{\partial x_1}|x_0(x-x_{10})+\frac{\partial f}{\partial x_2}|x_0(x_2-x_{20})$$

$$+\cdots+\frac{\partial f}{\partial x_7}|x_0(x_7-x_{70})+R_n, \qquad \text{其中 } R_n \text{ 为余项}$$

这里 x_i-x_{i0} 有明显的物理意义，即 x_i-x_{i0} 是零件 i 的随机误差，有 $x_i-x_{i0}\sim(0,\delta_i^2)$, 对随机变量 $x_i\sim(x_i,\delta_i^2)$.

我们认为展开式忽略 2 阶小后能够很好地逼近 y 值，理由主要有：

1. 就本问题我们可以求出： $\delta_{\max}=\delta_6=\frac{20\times 10\%}{3}=0.667$, 可见 x_i 的方差不大， $i=1\cdots 7$;

2. 当 x_i 与 x_{i0} 相差很小时， $x_i\sim N(x_i,\delta_i^2)$, $P(|x_i-x_{i0}|<\epsilon$ 很大，则在 $(x_1x_2\cdots x_n)$ 的领域内，近似值已很接近真实值.

当 x_i 与 x_{i0} 相差很大时， $P(|x_i-x_{i0}|<\epsilon)$ 很小，可以忽略其对真实值的影响，

所以在这个近似计算下： $y-f(x_{10},x_{20},\sim x_{70})$ 是 x_i 的随机误差 x_i-x_{i0} 的线性组合函数， $x_i-x_{i0}\sim N(0,\delta_i^2)$ 且 x_i-x_{i0} 互不相关.

定理 2 个独立正态随机变量 $\xi_k\sim N(\mu_k,\delta_k^2),(k=1,2)$ 的任意线性组合 $\sum_{i=1}^2\alpha_i\xi_i$ 为正态变量且 $\sum_{i=1}^2\alpha_i\xi_i\sim N(\sum_{i=1}^2\alpha_i^2,\mu_i^2,\sum_{i=1}^2\alpha_i\delta_i^2)$.

由此我们知道 $y-f(x_{10},\cdots,x_{70})$ 服从正态分布，由定理 $y=f(x_{10},\cdots,x_{70})+\sum\limits_{i=1}^{7}\frac{\partial f}{\partial x_i}(x_i-x_{i0})$ 服从正态分布，且

$$\bar{y}=E(y)=f(x_{10},\cdots,x_{70}),\qquad \delta^2=D(y)=\sum_{i=1}^{7}\left(\frac{\partial f}{\partial x_i}\,|\,X=X_0\right)^2\delta_i^2$$

这比我们期望的结果还要好.

基于统计抽样和近似算法的结论，我们得到 $y\sim N(f(x_{10},\cdots,x_{70}),\ \sum\limits_{i=1}^{7}(\frac{\partial f}{\partial x_i}\,|\,X=X_0)^2\delta_i^2)$. 为了进一步肯定我们的假设是正确的，我们利用样本数据作了 χ^2 拟合优度检验，在这里我们仅记录下检验结果 χ^2 （计算值）$=13.2410<\chi^2_{(0.05)}(11-2-1)=15.507$.

即 $y\sim N(f(x_{10},\cdots,x_{70}),\sum_{i=1}^{7}\left(\frac{\partial f}{\partial x_i}\,|\,X=X_0\right)$ 是可接受的.

（二） 模型及分析：

零件参数的优化设计：

目标函数： $\min(C_1+C_2)$, s.t. : $X_1\leq X\leq X_2$,

一件产品成本费用： $C_1\sum\limits_{i=1}^{7}\sum\limits_{j=1}^{3}a_{ij}b_{ij}$,

一件产品质量损失费用的期望值： $C_2=P_1\times1000+9000\times P_2$,

产品是次品的概率： $P_1=\frac{1}{\sqrt{2\pi}\delta}\left[\int_{y_0-0.3}^{y_0-0.1}e^{\frac{-(1-y)^2}{2\delta^2}}dt+\int_{y_0+0.1}^{y_0-0.3}e^{\frac{-(1-y)^2}{2\delta^2}}dt\right]$

产品的废品的概率： $P_2=\frac{1}{\sqrt{2\pi}\delta}\left[\int_{-\infty}^{y_0-0.3}e^{\frac{-(1-y)^2}{2\delta^2}}dt+\int_{y_0+0.3}^{-\infty}e^{\frac{-(1-y)^2}{2\delta^2}}dt\right]$,

$\delta_i=\left(\sum_{j=1}^{3}x_iB_i\times D_i'\right)/3$, $\delta^2=\sum_{i=1}^{7}\left[\left(\frac{\partial f}{\partial x_i}\right)\ \Big|_{x_0}\delta_0\right]^2$, $\overline{Y}=f(x_1,x_2,\cdots,x_7)$.

四、模型的求解

1） 原设计的生产费用：

由原设计知： $x=(0.1,0.3,0.1,0.1,1.5,16,0.75)$

$$B=\begin{bmatrix}0&1&0\\1&0&0\\1&0&0\\1&0&0\\1&0&0\\0&1&0\end{bmatrix}$$

即得：

$$\begin{aligned}&\bar{y}=1.725589262, &&\delta=0.1103719239,\\ &C_1=200, &&C_2=2874.79,\end{aligned}$$

生产总费用

$$\text{value}=(2874.79+200)\times1000=3,074,790(\text{元})$$

2） 优化模型的求解

目标函数可能存在局部最优解，这使我们寻找全局最优解变得很困难，为了避免搜索过程因陷于局部最优解而无法自拔，提高全局最优解的可信度，我们在这里采用模拟退火算法，其算法如下：

Step 1 给定 $X^0 \in R^n, T_0>0, \beta>0$, 给定产生随机向量的概率密度函数和控制温度下降过程的温度更新函数，给定常数 $\beta>0$, 计算 $f(x^0)$, 置 $X^0=x^0, X^0_{\min}=x_0, f_{\min}=f(x^0), k=0$;

Step 2 根据给定的概率密度函数产生一个随机向量 Z^k, 利用当前迭代点 X^k 和随机向量 Z^k 产生一个新的试探点 Y^k. 即 $Y^k=X^k+Z^k$, 计算 $f(Y^k)$;

Step 3 产生一个在 $(0,1)$ 上均匀分布的随机数 η, 计算在给定当前迭代点 X^k 和温度 T_k 下接受试探点 Y^k 的概率 $P_a[Y^k \mid X^k, T_k)$ 即

$$P_a\left[Y^k \mid X^k, T_k\right] = \min\left\{1, \exp\left(\frac{f(X^k)-f(Y^k)}{\beta T^k}\right)\right\}$$

如果 $\eta \le (P_a[Y^k \mid X^k, T_k)$, 则置 $X^{k+1}=Y^k, f(X^{k+1})=f(Y^k)$;

否则置 $X^{k+1}=X^k, f(X^{k+1})=f(X^k)$.

Step 4 如果 $f(X^{k+1})<f_{\min}$ 则置 $X_{\min}=X^{k+1}, f_{\min}=f(X^{k+1})$;

Step 5 如果迭代终止条件满足，则算法结束，$X_{\min}$ 就作为近似的全局最优解，$f_{\min}$ 为相应的最优值；否则继续 Step 6.

Step 6 根据给定的温度更新函数产生一个新的温度 T_{k+1}, 置 $k=k+1$ 转至 Step 2.

考虑到容差的组合，我们第一个想法就是对容差进行穷举，然后利用退火法来搜索最优解，但上机时，计算机内存不够，只能计算少数几组值；并且将均匀分布的随机数发生器换以不同的种子后，X 的“最优解”的波动较大，即在我们现在计算机上此法不合适，于是想到一个折中的算法.

第一步 我们确定零件的容差，然后利用退火算法多次循环计算总费用值，取最小值，记为 f_1^*, 以及对应的最优解 X_1^*, 转入第二步.

第二步 固定 X_{k-1}^*, 将零件的容差的所有组合进行穷举，计算得此时对应的地最优容差组合记为 B_k^*, 产生新 f^*, 最小值记为 f_k^*.

第三步 若 $f_{k+1}^*-f_k<\varepsilon$ (ε 是任意事先指定的常数), 终止，否则转入第一步.

利用上面的思想通过编程我们最后得到：

一批零件的总生产费用的最优解为 $434.6812*1000=434681.2$ 元. 容差向量为 [2,2,2,1,1,2,2]

五、模型结果的分析

在分析我们得到的最优解和最优值时，注意到在我们建模时是用正态分布函数代替经验分布函数的，在这里会产生误差.

我们在这个最优值上，进行随机抽样实验，共抽样 10 次，每次抽样 500 个随机数据计算每次抽样的质量损失费，然后求平均值算出在此点的平均质量损失费为 149.0, 即总生产成本为 $149.6+275=424.6$, 则实际成本较之可能还要小. 可见我们的最优值别较大.

这个误差一般来说是由于对经验分布函数进行估计所带来的，如果允许一定的误差，我们这种以模型的简练换精度还是值得的.

六、模型优缺点和改进方向（略）

七、附　　录（略）

参考文献

[1] 叶其孝，数学模型.
[2] 系统工程理论与实践，1997, 5.
[3] 耿素云，张立昂，概率统计，北京大学出版社，北京.
[4] 概率统计与随机过程，武汉大学出版社，武汉.
[5] 叶其孝主编，数学建模教育与国际数学建模竞赛.